

AD-A110 447

WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER  
AN ANALYSIS OF TRANSFORMATIONS REVISITED, REBUTTED.(U)  
OCT 81 8 E BOX, D R COX  
MRC-TSR-2288

F/8 12/1

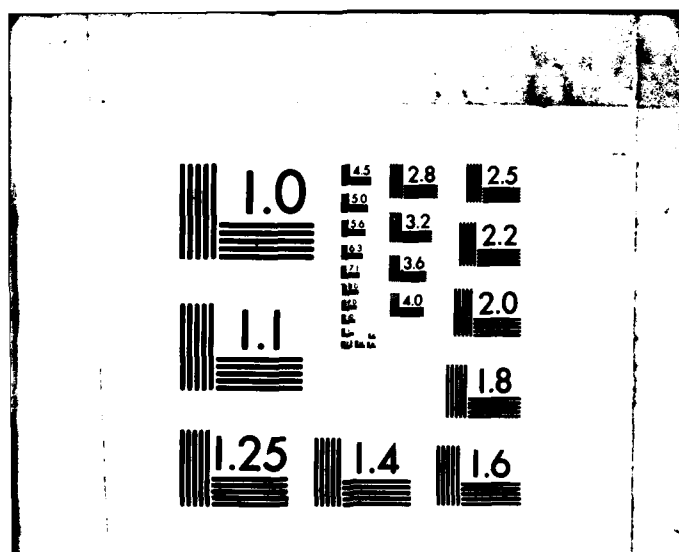
DAA029-80-C-0041

UNCLASSIFIED

1-1  
A10440



END  
DATE  
FILMED  
3 82  
DTIC



**LEVEL**

2

MRC Technical Summary Report #2288

AN ANALYSIS OF TRANSFORMATIONS  
REVISITED, REBUTTED

G. E. P. Cox and D. R. Cox

DTIC  
SELECTED  
FEB 4 1982

AD A110447

Mathematics Research Center  
University of Wisconsin-Madison  
610 Walnut Street  
Madison, Wisconsin 53706

October 1981

(Received Sept. 17, 1981)

Approved for public release  
Distribution unlimited

Sponsored by

U. S. Army Research Office  
P.O. Box 12211  
Research Triangle Park  
North Carolina 27709

82 02 03 077

DTIC FILE COPY

UNIVERSITY OF WISCONSIN-MADISON  
MATHEMATICS RESEARCH CENTER

AN ANALYSIS OF TRANSFORMATIONS REVISITED, REBUTTED

G. E. P. Box<sup>\*</sup> and D. R. Cox<sup>†</sup>

Technical Summary Report #2288

October 1981

ABSTRACT

In a recent paper, Bickel and Doksum (1981) argue that the performance of a method for estimating transformations due to Box and Cox is "unstable" because estimates of the transformation parameter on the one hand, and of the remaining parameters on the other, can be highly correlated. In this note it is argued that while criticisms are qualitatively obvious, they are scientifically irrelevant.

(pi)

AMS (MOS) Subject Classification: 62-07

Key Words: Transformation, Box-Cox

Work Unit Number 4 (Statistics and Probability)

---

<sup>\*</sup>University of Wisconsin, Madison, Wisconsin, U. S. A.

<sup>†</sup>Imperial College of Science and Technology, London, U. K.

---

<sup>\*</sup>Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

## SIGNIFICANCE AND EXPLANATION

→ In a ~~well-known~~ paper written in 1964, Box and Cox described a method for estimating transformations and showed how in suitable cases valuable increases in simplicity and efficiency were possible. Since that time, this technique has enjoyed wide practical use and considerable success. However, a recent theoretical paper by Bickel and Doksum (1981) seems to suggest that serious dangers are associated with the employment of this method, and speaks of "instability" and "cost" of estimation of the transformation. These difficulties seem to be associated with

(1) examples which common sense would rule out, namely situations where the effect of transformation on the data is almost linear, so that it is a matter of indifference which transformation is used;

(2) the idea that it makes sense to state conclusions in terms of a number measured on an arbitrary scale;

(3) failure to take proper account of the Jacobian of the transformation.

p11

Accession For	
NTIS GRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	



The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the authors of this report.

# AN ANALYSIS OF TRANSFORMATIONS REVISITED, REBUTTED

G. E. P. Box<sup>\*</sup> and D. R. Cox<sup>†</sup>

Transformation has long been a powerful tool in developing parsimonious representations and interpretations of data. In 1964 we examined the formal estimation of a suitable transformation. In particular suppose that a response  $y$  is transformed to  $y^{(\lambda)}$ , where

$$y^{(\lambda)} = \begin{cases} (y^\lambda - 1)/\lambda & (\lambda \neq 0) \\ \log y & (\lambda = 0) \end{cases}$$

and that we assume provisionally that for some unknown  $\lambda$ , the vector

$\underline{y}^{(\lambda)} = (y_1^{(\lambda)}, \dots, y_n^{(\lambda)})$  of  $n$  transformed observations satisfies a linear model

$$E(\underline{y}^{(\lambda)}) = \underline{X}\underline{\theta}$$

where  $\underline{\theta}$  is unknown, the errors being independently normally distributed with zero mean and constant variance  $\sigma^2$ . Estimation of  $\lambda$ ,  $\underline{\theta}$ , and  $\sigma^2$  can be by Bayesian or maximum likelihood methods.

Bickel and Doksum (1981) in a recent technically very impressive paper have in particular studied the joint estimation of  $\lambda$  and  $\underline{\theta}$ , examining

---

<sup>\*</sup>University of Wisconsin, Madison, Wisconsin, U. S. A.

<sup>†</sup>Imperial College of Science and Technology, London, U. K.

---

<sup>\*</sup>Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

consistency and asymptotic variances. They report that the cost of not knowing  $\lambda$  and having to estimate it, can be severe; that "...the performance of all Box-Cox type procedures is unstable and highly dependent on the parameters of the model in structured models with small to moderate error variances." That is, the estimates  $\hat{\lambda}$  and  $\hat{\theta}$  can be highly correlated, so that the marginal variances of the  $\hat{\theta}$ 's can be inflated by large factors over the conditional variances for fixed  $\lambda$ .

It seems to us that this general conclusion is qualitatively obvious and at the same time scientifically irrelevant.

To illustrate first the obviousness, take as a simple example the comparison of two groups of modest size, the observations  $y$  in group 1 being near 995 and those in group 2 being near 1005, the scatters within the two groups being roughly normal with standard deviations close to unity. A parameter  $\theta$  representing the difference between groups on the  $y$ -scale is quite precisely estimated to be about 10  $y$ -units. Suppose that the possibility of transformation were contemplated. For a very wide range of  $\lambda$  the function  $y^{(\lambda)}$  is very nearly linear in  $y$  over the span of the data, and, in particular, unless the sample sizes were very large indeed, it would be quite impossible to distinguish from the data whether  $y$  or  $y^{-1}$  gave better fit to the standard normal assumptions: if the parameter  $\theta$  were to refer to a difference on the  $y^{-1}$  scale it is quite precisely estimated to be near  $-10^{-5} y^{-1}$ -units (or  $10^{-5} y^{(-1)}$ -units, where  $y^{(-1)} = (1/y - 1)/(-1)$ ). Thus if the target parameter  $\theta$  is defined in terms of unknown  $\lambda$  in such a case as this, where  $\lambda$  is poorly determined, the numerical value of  $\theta$  (in units of  $y^{\lambda}$  or  $y^{(\lambda)}$ ) could be virtually anything.

As to the scientific implications of this, how can it be sensible scientifically to state a conclusion as a number measured on an unknown

scale? Surely to know that some effect has magnitude 10 units is without content unless one knows the scale and units in which the effect is defined. To say in the above idealized example that  $\theta$ , defining the difference between groups, is ill determined, because the data establish a wide range of functions as virtually equivalent, seems to be very misleading.

There is, of course, no dispute with Bickel and Doksum over mathematics: the issue is one of scientific relevance. As with any procedure it is necessary to use some common sense in estimating transformations, and in particular (see for example Box et al (1978), p. 241) not to expect this to be possible or relevant when for the particular data and the particular class of transformations in mind the transformation is essentially linear.

Of course the gross correlation effects would be avoided if, following our paper, the investigation had been conducted in terms of

$$z^{(\lambda)} = \frac{(y^\lambda - 1)/(\lambda \dot{y}^{(\lambda-1)})}{\dot{y} \log y} \quad \begin{array}{l} (\lambda \neq 0) \\ (\lambda = 0) \end{array}$$

which takes account of the Jacobian of the transformation. (For the above examples the differences in means for both  $z^{(1)}$  and  $z^{(-1)}$  would then have been very nearly ten units.) However some question of scientific relevance would still remain.

There are numerous aspects of transformations that merit further study. These include in particular the further development of simple ways of assessing transformation potential. That is, of providing some more formal measure of the ability of particular data to provide useful information about a class of transformations.



## REFERENCES

- Bickel, P.J. and Doksum, K.A. (1981). An analysis of transformations revisited. J. Amer. Statist. Assoc. 76, 296-311.
- Box, G.E.P. and Cox, D.R. (1964). An analysis of transformations. J. R. Statist. Soc. Series B 26, 211-252.
- Box, G.E.P., Hunter, W.G., and Hunter, J.S. (1978). Statistics for Experimenters. New York: John Wiley & Sons, Inc.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TSR #2288	2. GOVT ACCESSION NO. AD-A110 447	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle)  AN ANALYSIS OF TRANSFORMATIONS REVISITED, REBUTTED		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s)  G. E. P. Box and D. R. Cox		8. CONTRACT OR GRANT NUMBER(s)  DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of 610 Walnut Street Wisconsin Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS  4 (Statistics and Probability)
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P. O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE October 1981
		13. NUMBER OF PAGES 4
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)  UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  Transformation, Box-Cox		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  In a recent paper, Bickel and Doksum (1981) argue that the performance of a method for estimating transformations due to Box and Cox is "unstable" because estimates of the transformation parameter on the one hand, and of the remaining parameters on the other, can be highly correlated. In this note it is argued that while criticisms are qualitatively obvious, they are scientifically irrelevant.		